

**Use of secreted protein products for preventing and treating  
pancreatic diseases and/or obesity and/or metabolic syndrome**

**Description**

5 This invention relates to the use of secreted SF1-SF8 proteins, to the use of polynucleotides encoding these, and to the use of effectors/modulators thereof in the diagnosis, study, prevention, and treatment of pancreatic diseases (e.g. diabetes mellitus), obesity and/or metabolic syndrome and to the use in regeneration of tissues such as pancreatic tissues and others.

10 Many human proteins serve as pharmaceutically active compounds. Several classes of human proteins that serve as such active compounds include hormones, cytokines, cell growth factors, and cell differentiation factors. Most proteins that can be used as a pharmaceutically active compound fall within the family of secreted proteins. Secreted proteins are generally produced within cells at rough endoplasmic reticulum, are then exported to the golgi  
15 complex, and then move to secretory vesicles or granules, where they are secreted to the exterior of the cell via exocytosis. Examples for commercially used secreted proteins are human insulin, thrombolytic agents, interferons, interleukins, colony stimulating factors, human growth hormone, transforming growth factor beta, tissue plasminogen activator, erythropoietin, and various  
20 other proteins. Receptors of secreted proteins, which are membrane-bound proteins, also have potential as therapeutic or diagnostic agents. It is, therefore, important for developing new pharmaceutical compounds to identify secreted proteins that can be tested for activity in a variety of animal models. Thus, in light of the pervasive role of secreted proteins in human physiology, a  
25 need exists for identifying and characterizing novel functions for human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical diseases, disorders, and/or conditions by using secreted proteins or the genes that encode them.

30 The pancreas is an essential organ possessing both an exocrine function involved in the delivery of enzymes into the digestive tract and an endocrine function by which various hormones are secreted into the blood stream. The exocrine function is assured by acinar and centroacinar cells that produce various digestive enzymes and intercalated ducts that transport these enzymes

- 2 -

in alkaline solution to the duodenum. The functional unit of the endocrine pancreas is the islet of Langerhans. Islets are scattered throughout the exocrine portion of the pancreas and are composed of four cell types: alpha-, beta-, delta- and PP-cells, reviewed for example in Kim S.K. and Hebrok M., (2001) *Genes Dev.* 15: 111-127. Beta-cells produce insulin, represent the majority of the endocrine cells and form the core of the islets, while alpha-cells secrete glucagon and are located in the periphery. Delta-cells and PP-cells are less numerous and secrete somatostatin and pancreatic polypeptide, respectively.

Early pancreatic development has been well studied in different species, including chicken, zebrafish, and mice (for a detailed review, see Kim & Hebrok, 2001, *supra*). The pancreas develops from distinct dorsal and ventral anlagen. Pancreas development requires specification of the pancreas anlage along both anterior-posterior and dorsal-ventral axes. A number of transcription factors, which are critical for proper pancreatic development have been identified (see Kim & Hebrok, 2001, *supra*; Wilson M.E. et al., (2003) *Mech Dev.* 120: 65-80).

In postnatal/adult humans, the acinar and ductal cells retain a significant proliferative capacity that can ensure cell renewal and growth, whereas the islet cells become mostly mitotically inactive. This is in contrast to rodents where beta-cell replication is an important mechanism in the generation of new beta cells. It has been suggested, that during embryonic development, pancreatic islets of Langerhans originate from differentiating duct cells or other cells with epithelial morphology (Bonner-Weir S. and Sharma A., (2002) *J Pathol.* 197: 519-526; Gu G. et al., (2003) *Mech Dev.* 120: 35-43). In adult humans, new beta cells arise in the vicinity of ducts (Butler A.E. et al., (2003) *Diabetes* 52: 102-110; Bouwens L. and Pipeleers D.G., (1998) *Diabetologia* 41: 629-633). However, also an intra-islet location or an origin in the bone marrow has been suggested for precursor cells of adult beta cells (Zulewski H. et al., (2001) *Diabetes* 50: 521-533; Janus A. et al., (2003) *J Clin Invest.* 111: 843-850). Pancreatic islet growth is dynamic and responds to changes in insulin demand, such as during pregnancy or during the increase in body mass occurring during childhood. In adults, there is a good correlation between body mass and islet mass (Yoon K.H. et al., (2003) *J Clin Endocrinol Metab.* 88: 2300-2308).

Pancreatic beta-cells secrete insulin, which is stimulated by high blood glucose levels. Insulin amongst other hormones plays a key role in the regulation of the fuel metabolism. Insulin leads to the storage of glycogen and triglycerides and to the synthesis of proteins. The entry of glucose into muscles and adipose cells is stimulated by insulin. In patients who suffer from diabetes mellitus type I or LADA (latent autoimmune diabetes in adults (Pozzilli & Di Mario, 2001, Diabetes Care. 8:1460-67) beta-cells are being destroyed due to autoimmune attack. The amount of insulin produced by the remaining pancreatic islet cells is too low, resulting in elevated blood glucose levels (hyperglycemia). In diabetes type II, liver and muscle cells lose their ability to respond to normal blood insulin levels (insulin resistance). High blood glucose levels (and also high blood lipid levels) lead to an impairment of beta-cell function and to an increase in beta-cell apoptosis. It is interesting to note that the rate of beta-cell neogenesis does not appear to change in type II diabetics (Butler et al., 2003 supra), thus causing a reduction in total beta-cell mass over time. Eventually the application of exogenous insulin becomes necessary in type II diabetics.

Improving metabolic parameters such as blood sugar and blood lipid levels (e.g. through dietary changes, exercise, medication or combinations thereof) before beta cell mass has fallen below a critical threshold leads to a relatively rapid restoration of beta cell function. However, after such a treatment the pancreatic endocrine function would remain impaired due to the only slightly increased regeneration rate.

In type I diabetics, the lifespan of pancreatic islets is dramatically shortened due to autoimmune destruction. Treatments have been devised which modulate the immune system and may be able to stop or strongly reduce islet destruction (Raz I. et al., (2001) Lancet 358: 1749-1753; Chatenoud L. et al., (2003) Nat Rev Immunol. 3: 123-132). However, due to the relatively slow regeneration of human beta cells such treatments could only be fully successful at improving the diabetic condition if they are combined with an agent which can stimulate beta cell regeneration.

Thus, both for type I and type II diabetes (early and late stages) there is a need to find novel agents which stimulate beta cell regeneration.

Diabetes is a very disabling disease, because today's common anti-diabetic drugs do not control blood sugar levels well enough to completely prevent the occurrence of high and low blood sugar levels. Out of range blood sugar levels are toxic and cause long-term complications like for example renopathy,  
5 retinopathy, neuropathy, and peripheral vascular disease. There are also a host of related conditions, such as obesity, hypertension, heart disease, and hyperlipidemia, for which persons with diabetes are at substantially increased risk.

10 Apart from the impaired quality of life for the patients, the treatment of diabetes and its long-term complications presents an enormous financial burden to our healthcare systems with rising tendency. Thus, for the prevention or treatment of, type I and type II diabetes as well as for latent autoimmune diabetes in adults (LADA) there is a strong need in the art to  
15 identify factors that induce regeneration of pancreatic insulin producing beta-cells. These factors could restore normal function of the endocrine pancreas once its function is impaired or event could prevent the development or progression of diabetes type I, diabetes type II, or LADA.

20 Obesity is one of the most prevalent metabolic disorders in the world. It is still a poorly understood human disease that becomes as a major health problem more and more relevant for western society. Obesity is defined as a body weight more than 20% in excess of the ideal body weight, frequently resulting in a significant impairment of health. Obesity may be measured by body mass  
25 index, an indicator of adiposity or fatness. Further parameters for defining obesity are waist circumferences, skinfold thickness and bioimpedance. It is associated with an increased risk for cardiovascular disease, hypertension, diabetes mellitus type II, hyperlipidaemia and an increased mortality rate. Obesity is influenced by genetic, metabolic, biochemical, psychological, and  
30 behavioral factors and can be caused by different reasons such as non-insulin dependent diabetes, increase in triglycerides, increase in carbohydrate bound energy and low energy expenditure (Kopelman P.G., (2000) Nature 404: 635-643).

35 The concept of 'metabolic syndrome' (syndrome x, insulin-resistance

syndrome, deadly quartet) was first described 1966 by Camus and reintroduced 1988 by Reaven (Camus J.P., (1966) *Rev Rhum Mal Osteoartic* 33: 10-14; Reaven G.M., (1988), *Diabetes* 37: 1595-1607). Today, metabolic syndrome is commonly defined as clustering of cardiovascular risk factors like hypertension, abdominal obesity, high blood levels of triglycerides and fasting glucose as well as low blood levels of HDL cholesterol. Insulin resistance greatly increases the risk of developing the metabolic syndrome (Reaven G., (2002) *Circulation* 106: 286-288). The metabolic syndrome often precedes the development of type II diabetes and cardiovascular disease (Lakka H.M. et al., (2002) *JAMA* 288: 2709-2716). The control of blood lipid levels and blood glucose levels is essential for the treatment of the metabolic syndrome (see, for example, Santomauro A.T. et al., (1999) *Diabetes* 48: 1836-1841).

The molecular factors regulating food intake and body weight balance are incompletely understood. Even if several candidate genes have been described which are supposed to influence the homeostatic system(s) that regulate body mass/weight, like leptin or the peroxisome proliferator-activated receptor-gamma co-activator, the distinct molecular mechanisms and/or molecules influencing obesity or body weight/body mass regulations are not known.

In this invention, we disclose secreted factors referred to as SF-1-SF8, which are involved in pancreas development, regeneration, and in the regulation of energy homeostasis.

No functional data are available in the scientific prior art for the secreted protein referred to as SF1. However, the SF1 protein is structurally related to chitinases. The fly homologue of SF1 lost its enzymatic activity. It is expressed in the fat body of the *Drosophila* embryo and larva, and might act as a growth factor (Kawamura K. et al., (1999) *Development* 126: 211-219).

No functional data are available in the scientific prior art for the secreted protein referred to as SF2. Expressed sequence tags (ESTs) of the SF2 mRNA are present, e.g. in pancreas and ES cells. The SF2 gene is highly conserved in mouse, rat, and humans.

The extracellular matrix protein SF3 is involved in the growth and guidance

of axons in the spinal cord and the PNS (Klar A. et al., (1992) Cell. 69: 95-110). The human SF3 gene is located at the region of chromosome 11 (between markers D11S926 and D11S928); deletion of this region causes nesidioblastosis (hyperinsulinemia) in humans. SF3 was shown to affect the biological activities of vascular endothelial cells by inhibiting angiogenesis (Terai Y. et al., (2001) J Cell Physiol. 188: 394-402). Furthermore SF3 was shown to function as a contact-repellent molecule for embryonic motor neurons (Tzarfati-Majar V. et al., (2001) Proc Natl Acad Sci U S A. 98: 4722-4727).

SF4 belongs to a structural family of neuropeptides that are widely distributed in the body and function as neurotransmitters and neuromodulators (Goto T. and Tanaka T., (2002) Microsc Res Tech. 58: 91-97). A regulatory role for SF4 in placental physiology and pre-eclampsia was described (Page N.M. et al., (2001) Regul Pept. 98: 97-104).

SF5 is a glycoprotein involved in cell anchorage and integrin signaling. SF5 is secreted from mammary gland and expressed in outer mesenchymal stromal region of the mouse fetal gonad (Oshima K. et al., (1999) Biochem Biophys Res Commun. 254: 522-528; Kanai Y. et al., (2000) Mech Dev. 96: 223-227). A possible role of SF5 in the membrane vesicle secretion and exocytosis of endocytic multivesicular bodies was suggested (Oshima K. et al., (2002) Eur J Biochem. 269: 1209-1218).

SF6 is a secreted glycoprotein expressed in CNS, PNS, and subsets of mesenchymal cells (Nelson B.R. et al., (2002) Gene Expr Patterns. 2: 7-15). SF6 contains a calcium ion binding domain.

SF7 is a serine protease inhibitor, that regulates plasma levels of glucocorticoids. Moreover SF7 is the main carrier of glucocorticoids in mammals. The temporal and spatial changes in the localization of SF7 and its mRNA in the fetus are suggested to influence the effects of steroid hormones on developing tissues (Scrocchi L.A. et al., (1993) Endocrinology. 132: 903-909). Furthermore it was suggested that during severe stress, SF7 might be important in regulating the amount of cortisol reaching target tissues (Garrel D.R. (1996) Horm Res. 45: 245-251).

SF8 is a cyclosporin A binding protein, which is expressed in a restricted subset of tissues including mouse ovary, testis, bone marrow, and kidney, and human kidney, pancreas, skeletal muscle, heart, lung, and liver (Friedman J. et al., (1994) Am J Pathol. 144: 1247-1256; Schneider H. et al., (1994) Biochemistry. 33: 8218-8224). SF8 is involved in degradation of the genome during apoptosis (Montague J.W. et al., (1997) J Biol Chem. 272: 6677-6684).

Accordingly, the present invention relates to secreted proteins with novel functions in the human metabolism, regeneration, and pancreatic developmental processes.

The present invention discloses specific genes and proteins encoded thereby and effectors/modulators thereof involved in the regulation of pancreatic function and metabolism, especially in pancreas diseases such as diabetes mellitus, e.g. insulin dependent diabetes mellitus and/or non insulin dependent diabetes mellitus, and/or LADA and/or metabolic syndrome, obesity, and/or related disorders such as coronary heart disease, eating disorder, cachexia, hypertension, hypercholesterolemia (dyslipidemia), liver fibrosis, and/or gallstones. Further, the present invention discloses specific genes and proteins encoded thereby and effectors/modulators thereof involved in the regeneration of pancreatic cells or tissues, e.g. cells having exocrine functions such as acinar cells, centroacinar cells and/or ductal cells and/or cells having endocrine functions, particularly cells in Langerhans islets such as alpha-, beta-, delta- and/or PP-cells, more particularly beta-cells.

In this invention, we used a screen for secreted factors expressed in developing mammalian (mouse) pancreas, as described in more detail in the Examples section (see Example 1). This screen identified SF1-SF8 as secreted factors expressed in developing mouse pancreas. The present invention describes mammalian SF1-SF8 proteins and the polynucleotides encoding these, in particular human SF1-SF8, as being involved in the conditions and processes mentioned above.

The present invention relates to SF1, SF2, SF3, SF4, SF5, SF6, SF7 and/or SF8 (SF1-SF8) polynucleotides encoding polypeptides with novel functions in the development and regeneration of pancreatic tissues and thus in

mammalian pancreatic diseases (e.g. diabetes), and also in body-weight regulation, energy homeostasis, and obesity, fragments of said polynucleotides, polypeptides encoded by said polynucleotides or fragments thereof. The invention also relates to vectors, host cells, and recombinant methods for producing the polypeptides and polynucleotides of the invention. The invention also relates to effectors/modulators of SF1-SF8 polynucleotides and/or polypeptides, e.g. antibodies, biologically active nucleic acids, such as antisense molecules, RNAi molecules or ribozymes, aptamers, peptides or low-molecular weight organic compounds recognizing said polynucleotides or polypeptides.

SF1-SF8 homologous proteins and nucleic acid molecules coding therefore are obtainable from vertebrate species. Particularly preferred are nucleic acids encoding the human SF1-SF8 protein and variants thereof. The invention particularly relates to a nucleic acid molecule encoding a polypeptide contributing to regulating the energy homeostasis and the mammalian metabolism, wherein said nucleic acid molecule comprises

- (a) the nucleotide sequence of human SF1-SF8 and/or a sequence complementary thereto,
- (b) a nucleotide sequence which hybridizes at 50°C in a solution containing 1 x SSC and 0.1% SDS to a sequence of (a),
- (c) a sequence corresponding to the sequences of (a) or (b) within the degeneration of the genetic code,
- (d) a sequence which encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and up to 99,6% identical to the amino acid sequences of the human SF1-SF8 proteins,
- (e) a sequence which differs from the nucleic acid molecule of (a) to (d) by mutation and wherein said mutation causes an alteration, deletion, duplication and/or premature stop in the encoded polypeptide or
- (f) a partial sequence of any of the nucleotide sequences of (a) to (e) having a length of 15-25 bases, preferably 25-35 bases, more preferably 35-50 bases and most preferably at least 50 bases.

The function of the mammalian SF1-SF8 in mammalian metabolism was validated by analyzing the expression of the transcripts in different tissues and by analyzing the role in adipocyte differentiation (see Examples 3 and 4 for



more detail).

Expression profiling studies (see Examples for more detail) confirm the particular relevance of SF1-SF8 as regulators of energy metabolism in mammals.

We used mouse models of insulin resistance and/or diabetes, such as mice carrying gene knockouts in the leptin pathway (for example, ob/ob (leptin) or db (leptin receptor/ligand) mice) to study the expression of SF1-SF8. Such mice develop typical symptoms of diabetes, show hepatic lipid accumulation and frequently have increased plasma lipid levels (see Bruning J.C. et al., (1998) Mol. Cell. 2: 559-569). Expression of SF1-SF8 mRNA was also examined in susceptible wild type mice (for example, C57Bl/6) that show symptoms of diabetes, lipid accumulation, and high plasma lipid levels, if fed a high fat diet. Further, expression of SF4 mRNA was also examined in Non-Obese-Diabetic (NOD) mice.

Microarrays are analytical tools routinely used in bioanalysis. A microarray has molecules distributed over, and stably associated with, the surface of a solid support. The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, antibodies, or other chemical compounds on a substrate. Microarrays of polypeptides, polynucleotides, and/or antibodies have been developed and find use in a variety of applications, such as monitoring gene expression, drug discovery, gene sequencing, gene mapping, bacterial identification, and combinatorial chemistry. One area in particular in which microarrays find use is in gene expression analysis (see Example 4). Array technology can be used to explore the expression of a single polymorphic gene or the expression profile of a large number of related or unrelated genes. When the expression of a single gene is examined, arrays are employed to detect the expression of a specific gene or its variants. When an expression profile is examined, arrays provide a platform for identifying genes that are tissue specific, are affected by a substance being tested in a toxicology assay, are part of a signaling cascade, carry out housekeeping functions, or are specifically related to a particular genetic predisposition, condition, disease, or disorder.

Microarrays may be prepared, used, and analyzed using methods known in the

art (see for example, Brennan T.M., (1995) U.S. Patent No. 5,474,796; Schena M. et al., (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschwieler J.D. et al., (1995) PCT application WO 95/251116; Shalon T.D. and Brown P.O., (1995) PCT application WO 95/35505; Heller R.A. et al., (1997) Proc. Natl. Acad. Sci. USA 94: 2150-2155; Heller, M.J. and Tu E., (1997) U.S. Patent No. 5,605,662). Various types of microarrays are well known and thoroughly described in Schena M., ed. (1999; DNA Microarrays: A Practical Approach, Oxford University Press, London).

Oligonucleotides or longer fragments derived from any of the polynucleotides described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques, which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents, which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

As determined by microarray analysis, SF3 and SF5 show differential expression in human adipocytes. A strong up-regulation is observed concerning the expression of SF3 during the human primary adipocyte differentiation (see Fig. 9). A strong down-regulation is observed concerning the expression of SF5 during the human primary adipocyte differentiation (see Fig. 10A) and human SGBS cell differentiation (see Fig. 10B). The SF3 protein in preadipocytes has the potential to enhance adipocyte differentiation at a very early stage, and the SF5 protein in preadipocytes has the potential to inhibit adipocyte differentiation at a very early stage. Therefore, the SF3 and SF5 proteins might play an essential role in adipogenesis. The results are suggesting a role of SF3 and SF5 in the regulation in human metabolism, for example, as effectors/modulators (for example, inhibitors or enhancers) of adipogenesis. Thus, SF3 and SF5 are strong candidates for the manufacture

of pharmaceutical compositions and medicaments for the treatment of conditions related to human metabolism, such as diabetes, obesity, and/or metabolic syndrome.

5 Further, we show sectional in situ hybridizations of SF1 (see Example 5 and Fig. 11). The nucleic acid sequence encoding the mouse SF1 protein is expressed in the pancreas tissue (see Fig. 11). SF1 showed widespread expression in non-beta cell and non-duct cells. The distribution and shape of the expressing cells as well as co-staining with the exocrine marker amylase  
10 indicate that SF1 is expressed predominantly in exocrine cells in the developing pancreas.

The invention also encompasses novel use of polynucleotides that encode the proteins of the invention. Accordingly, any nucleic acid sequence, which  
15 encodes the amino acid sequences of the proteins of the invention, can be used to generate recombinant molecules that express the proteins of the invention. In a particular embodiment, the invention encompasses a nucleic acid encoding SF1-SF8. It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide  
20 sequences encoding the proteins, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. The invention contemplates each and every possible variation of nucleotide sequence that can be made by selecting combinations based on possible codon choices.

25 Also encompassed by the invention is the use of polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those of the polynucleotide encoding the proteins of the invention, under various conditions of stringency. Hybridization conditions are based on  
30 the melting temperature ( $T_m$ ) of the nucleic acid binding complex or probe, as taught in Wahl G.M. et al., (1987; Methods Enzymol. 152: 399-407) and Kimmel A.R. (1987; Methods Enzymol. 152: 507-511), and may be used at a defined stringency. Preferably, hybridization under stringent conditions means that after washing for 1 h with 1 x SSC and 0.1% SDS at 50°C, preferably at  
35 55°C, more preferably at 62°C and most preferably at 65°C, particularly for 1 h in 0.2 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C and most preferably at 65°C, a positive hybridization signal is observed.

Altered nucleic acid sequences encoding the proteins which are encompassed by the invention include deletions, insertions or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent protein.

5 The encoded proteins may also contain deletions, insertions or substitutions of amino acid residues, which produce a silent change and result in functionally equivalent proteins. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity  
10 of the protein is retained. Furthermore, the invention relates to peptide fragments of the proteins or derivatives thereof such as cyclic peptides, retro-inverso peptides or peptide mimetics having a length of at least 4, preferably at least 6 and up to 50 amino acids.

15 Also included within the scope of the present invention are alleles of the genes encoding the proteins of the invention. As used herein, an 'allele' or 'allelic sequence' is an alternative form of the gene, which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or polypeptides whose structures or function may or may not be  
20 altered. Any given gene may have none, one or many allelic forms. Common mutational changes, which give rise to alleles, are generally ascribed to natural deletions, additions or substitutions of nucleotides. Each of these types of changes may occur alone or in combination with the others, one or more times  
25 in a given sequence.

The nucleic acid sequences encoding SF1-SF8 may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements.

30 In order to express a biologically active protein, the nucleotide sequences encoding the proteins or functional equivalents, may be inserted into appropriate expression vectors, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods, which are well known to those skilled in the art, may be used to  
35 construct expression vectors containing sequences encoding the proteins and the appropriate transcriptional and translational control elements. Regulatory

elements include for example a promoter, an initiation codon, a stop codon, a mRNA stability regulatory element, and a polyadenylation signal. Expression of a polynucleotide can be assured by (i) constitutive promoters such as the Cytomegalovirus (CMV) promoter/enhancer region, (ii) tissue specific promoters such as the insulin promoter (see, Soria B. et al., (2000), Diabetes 49: 157-162), SOX2 gene promoter (see Li M. et al., (1998) Curr. Biol. 8: 971-974), Msi-1 promoter (see Sakakibara S. and Okano H., (1997) J. Neuroscience 17: 8300-8312), alpha-cardia myosin heavy chain promoter or human atrial natriuretic factor promoter (Klug M.G. et al., (1996) J. Clin. Invest 98: 216-224; Wu J. et al., (1989) J. Biol. Chem. 264: 6472-6479) or (iii) inducible promoters such as the tetracycline inducible system. Expression vectors can also contain a selection agent or marker gene that confers antibiotic resistance such as the neomycin, hygromycin or puromycin resistance genes. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook J. et al., (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y. and Ausubel F.M. et al., (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y.

In a further embodiment of the invention, natural, modified or recombinant nucleic acid sequences encoding the proteins of the invention may be ligated to a heterologous sequence to encode a fusion protein.

A variety of expression vector/host systems, as known in the art, may be utilized to contain and express sequences encoding the proteins or fusion proteins. These include, but are not limited to, micro-organisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus, adenovirus, adeno-associated virus, lentivirus, retrovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or PBR322 plasmids); or animal cell systems.

The presence of polynucleotide sequences of the invention in a sample can be detected by DNA-DNA or DNA-RNA hybridization and/or amplification using

probes or portions or fragments of said polynucleotides. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences specific for the gene to detect transformants containing DNA or RNA encoding the corresponding protein. As used herein  
5 'oligonucleotides' or 'oligomers' refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or amplimer.

10 A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting polynucleotide sequences include oligo-labeling, nick translation, end-labeling  
15 of RNA probes, PCR amplification using a labeled nucleotide, or enzymatic synthesis. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, Mich.); Promega (Madison Wis.); and U.S. Biochemical Corp., (Cleveland, Ohio).

The presence of SF1-SF8 in a sample can be determined by immunological  
20 methods or activity measurement. A variety of protocols for detecting and measuring the expression of proteins, using either polyclonal or monoclonal antibodies specific for the protein or reagents for determining protein activity are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting  
25 (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on the protein is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.) and Maddox D.E.  
30 et al. (1983; J. Exp. Med. 158: 1211-1226).

Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent or chromogenic agents as well as substrates, co-factors, inhibitors, magnetic particles, and the like.

35

The nucleic acids encoding the proteins of the invention can be used to generate transgenic animal or site specific gene modifications in cell lines.

Transgenic animals may be made through homologous recombination, where the normal locus of the genes encoding the proteins of the invention is altered. Alternatively, a nucleic acid construct is randomly integrated into the genome. Vectors for stable integration include plasmids, retroviruses and other animal viruses, YACs, and the like. The modified cells or animal are useful in the study of the function and regulation of the proteins of the invention. For example, a series of small deletions and/or substitutions may be made in the genes that encode the proteins of the invention to determine the role of particular domains of the protein, functions in pancreatic differentiation, etc.

Specific constructs of interest include anti-sense molecules, which will block the expression of the proteins of the invention, or expression of dominant negative mutations. A detectable marker, such as for example lac-Z, may be introduced in the locus of the genes of the invention, where upregulation of expression of the genes of the invention will result in an easily detected change in phenotype.

One may also provide for expression of the genes of the invention or variants thereof in cells or tissues where it is not normally expressed or at abnormal times of development. In addition, by providing expression of the proteins of the invention in cells in which they are not normally produced, one can induce changes in cell behavior.

DNA constructs for homologous recombination will comprise at least portions of the genes of the invention with the desired genetic modification, and will include regions of homology to the target locus. DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and/or negative selection are included. Methods for generating cells having targeted gene modifications through homologous recombination are known in the art. For non-human embryonic stem (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, e.g. mouse, rat, guinea pig etc. Such cells are grown on an appropriate fibroblast-feeder layer or grown in presence of leukemia inhibiting factor (LIF).

The data disclosed in this invention show that the SF1-SF8 nucleic acids and proteins and effector/modulator molecules thereof are useful in diagnostic and

therapeutic applications implicated, for example, but not limited to, pancreatic diseases (e.g. diabetes, such as insulin dependent diabetes mellitus and/or non insulin dependent diabetes mellitus and/or LADA), obesity, metabolic syndrome, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and/or gallstones. Further, the data show that the SF1-SF8 nucleic acids and proteins and effector/modulator molecules thereof are useful for the modulation, e.g. stimulation of pancreatic development, and/or for the regeneration of pancreatic cells or tissues, e.g. cells having exocrine functions such as acinar cells, centroacinar cells and/or ductal cells and/or cells having endocrine functions, particularly cells in Langerhans islets such as alpha-, beta-, delta- and/or PP-cells, more particularly beta-cells. Hence, diagnostic and therapeutic uses for the proteins of the invention nucleic acids and proteins of the invention are, for example but not limited to, the following: (i) tissue regeneration in vitro and in vivo (regeneration for all these tissues and cell types composing these tissues and cell types derived from these tissues), (ii) small molecule drug target, (iii) antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) diagnostic and/or prognostic marker, (v) protein therapy, (vi) gene therapy (gene delivery/gene ablation), and / or (vii) research tools.

For example, but not limited to, cDNAs encoding the proteins of the invention and particularly their human homologues may be useful in stimulating, enhancing or regulating the regeneration of tissues, and the proteins of the invention and particularly their human homologues may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from, for example, pancreatic diseases (e.g. diabetes), obesity, and/or metabolic syndrome as described above.

In one embodiment of the invention, administration of SF1-SF8 nucleic acids and proteins and/or effectors/modulators thereof in a pharmaceutical composition to a subject in need thereof, particularly a human patient leads to an at least partial regeneration of, for example, pancreas cells, e.g. endocrine cells and/or exocrine cells. The composition will then at least partially restore normal pancreatic function. In one example, these cells are beta cells of the islets or exocrine cells which will contribute to the improvement of a diabetic state. After administration of this composition on a short term or regular basis,



an increase in cell mass can be achieved. This effect upon the body reverses the condition of diabetes partially or completely. As the subject's blood sugar level improves, the dosage administered may be reduced in strength. In at least some cases further administration can be discontinued entirely and the subject continues to produce a normal amount of insulin without further treatment. The subject is thereby not only treated but cured entirely of a diabetic condition. However, even moderate improvements in cell mass can lead to a reduced requirement for exogenous insulin, improved glycemic control and a subsequent reduction in diabetic complications. In another example, other cells of the pancreas can be regenerated in vivo and/or in vitro to cure a certain condition. Beside diabetes, the compositions of the present invention will also have efficacy for treatment of patients with other pancreatic diseases such as pancreatic cancer, dysplasia, or pancreatitis.

The SF1-SF8 nucleic acids and proteins and effectors/modulators thereof are useful in diagnostic and therapeutic applications implicated in various embodiments as described below. For example, but not limited to, cDNAs encoding the proteins of the invention and particularly their human homologues may be useful in gene therapy, and the proteins of the invention and particularly their human homologues may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from, for example, pancreatic diseases (e.g. diabetes), obesity, and/or metabolic syndrome as described above.

The SF1-SF8 nucleic acids and proteins and effectors/modulators thereof may be administered either as a monotherapy or as a combination therapy with other pharmaceutical agents. For example, they may be administered together with other pharmaceutical agents suitable for the treatment or prevention of pancreatic diseases and/or obesity and/or metabolic syndrome. Further, they may be administered together with pharmaceutical agents which have an immunosuppressive activity, e.g. antibodies, polypeptides and/or peptidic or non-peptidic low molecular weight substances. Preferred examples of immunosuppressive agents are listed in the following Table 1.

Table 1: Exemplary agents for immune suppression

Names	Mechanism
2-amino-1,3-propanediol derivatives	Used for preventing or treating chronic rejection in a patient receiving an organ or tissue allo- or xenotransplant
2-amino-2[2-(4-octylphenyl)ethyl]propane-1,3-diol hydrochloride	Immunosuppression, from accelerated lymphocyte homing
40-O-(2-hydroxyethyl)-rapamycin, SDZ-RAD, Everolimus	Sirolimus (rapamycin) derivative, used for acute kidney rejection; reduces rejection and graft vasculopathy following heart transplantation by inhibiting cell proliferation
6-(3-dimethyl-aminopropionyl) forskolin	Immunosuppressing action useful also for treating autoimmune disease
6-mercaptopurine (6-MP)	Used to treat Crohn's disease, inflammatory bowel disease and for organ transplant therapy
A-420983	Lck-inhibitor
ABX-CBL (CBL-1)	Mouse monoclonal AB targeted against human T-cell, B cells, NK cells and monocytes, for treatment of steroid-resistant graft vs host diseases, potential use in treatment of inflammatory and autoimmune disorders
Alefacept (human LFA-3 IgG1 fusion protein)	Knocks out causative memory T-lymphocytes; Used to treat psoriasis, a T-cell mediated inflammatory disorder
Antisense ICAM-1 inhibitor (ISIS 2302), Enlimomab, BIRR1, Alicaforsen	Mouse monoclonal AB blocks white blood cell adhesion to T-cell surface molecule (ICAM-1r); treatment of kidney transplant rejection
Antithymocyte immunoglobulin (ATGAM)	Anti-human thymocyte, immunoglobulin; used in reversal of acute kidney transplant rejection and will likely be used off-label for transplant induction therapy
Azathioprine	Treatment of rheumatoid arthritis and prevention of kidney transplant rejection, and other autoimmune or inflammatory disorders such as inflammatory bowel disease
Baohuoside-1	Flavonoid; inhibits lymphocyte activation; Ma et al., Transplantation 78: 831-838, (2004)
basiliximab	Monoclonal AB that binds to receptor sites on T-cells, preventing activation by transplanted tissue (renal transplant)
BMS-279700	Lck-inhibitor
BTI-322	Mouse derived monoclonal AB targeted to CD2 receptor; used for prevention of first-time kidney rejection, and treatment of resistant rejection
Cladribine	Antimetabolite and immunosuppressive agent that is relatively selective for lymphocytes; used to treat lymphoid malignancies, e.g., hairy-cell leukemia
CP-690550	JAK-3 inhibitor
Cyclophosphamide (CTX)	Immunosuppressant for treatment of arthritis and other autoimmune disorders and cancers
Cyclosporine (cyclosporin A, cyclosporin)	11 amino acid cyclic peptide; blocks helper T-cell, immunosuppressant used in organ transplant therapy and other immune diseases
Daclizumab, HAT (Humanized Anti-Tac), SMART anti-Tac, anti-CD25, and humanized anti-IL2-receptor	Monoclonal AB inhibits binding of IL-2 to IL-2 receptor by binding to IL-2 receptor; suppresses T-cell activity against allografts (renal transplant)
Dexamethasone (Decadron, Dexone, Dexasone)	An adrenocorticoid, effective immunosuppressant in various disorders
DIAPEP-277	Immunomodulatory properties
Dipeptide Boronic Acid (DPBA)	Proteasome inhibitor; Wu et al., Transplantation 78: 360-366, (2004)
Docosahexaenoic acid (DHA)	Immunosuppressant that lowers the proportion of T cells expressing CD4 or CD8, blocks antigen recognition process; Taku et al., Journal of Agricultural and Food Chemistry 48: 1047, (2000)
efalizumab	T-cell modulator that target T-cells through interactions with adhesion molecules on endothelial cell surface, target migration of T-cells into the skin and target activation of T-cells; Used to treat Psoriasis

Efomycine M	Leukocyte adhesion inhibitor, Anti-inflammatory
FTY720 (oral myriocin derivative)	Alters lymphocyte infiltration into grafted tissues; used for prevention of organ rejection in kidney transplants
Glatiramer acetate (co-polymer-1)	Synthetic peptide copolymer; decoy that mimics structure of myelin so immune cells bind Copaxone instead of myelin; for multiple sclerosis
Glial fibrillary acidic protein (GFAP)	Possesses immunosuppressive activities in diabetic animal models; Winer et al., Nature Medicine 9: 198, (2003)
Gusperimus (15-deoxyspergualin)	Intravenous immunosuppressant; suppresses production of cytotoxic T-cells, neutrophils and macrophages
HLA-B2702 peptide	Human peptide, blocks action of NK cells and T-cell mediated toxicities, used for prevention of first kidney allograft rejection
hu1124 (anti-CD11a)	Humanized monoclonal antibody; targets CD11a receptor on surface of T cells to selectively inhibit immune system rejection of transplanted organs
hOKT31 $\gamma$ (Ala-Ala)	non Fc-binding humanized anti CD3 antibody
Infliximab	Monoclonal AB, binds and inactivates human TNF $\alpha$ ; used to treat Crohn's disease and rheumatoid arthritis
Interferon	Immunomodulatory properties
ISAbx247	Used to treat autoimmune diseases such as rheumatoid arthritis and psoriasis
isotretinoin	Immunosuppressant, reduces ability of T cells to proliferate in response to immune challenge. Vergelli et al., Immunopharmacology, 31:191, (1997)
L-683,742: also described as 31-desmethoxy-31-hydroxy-L-683,590	Treatment of autoimmune diseases, infectious diseases and/or prevention of organ transplant rejections
Leflunomide (ARAVA)	Antiinflammatory agent
Medi-500 (T10B9)	Intravenous monoclonal AB that targets human T-cells; treats acute kidney rejection and graft-vs-host disease
Medi-507	Intravenous humanized AB directed against CD2 T-cell; used to treat corticosteroidresistant graft vs host disease and prevention of kidney rejection
Methotrexate	Antimetabolite used to treat Crohn's disease, severe psoriasis, and adult rheumatoid arthritis (and as an anti-cancer drug)
Mitoxantrone	Antiproliferative effect on cellular immune system including T-cells, B-cells and macrophages; used to treat hormone-refractory prostate cancer, acute myelogenous leukemia and multiple sclerosis
mycophenolate mofetil	Proliferation of T and B lymphocytes by blocking the synthesis of purine nucleotides; used in organ transplant therapy and inflammatory bowel disease
OKT4A	Mouse monoclonal AB targeted against human CD4 T cell; used for prevention of kidney transplant rejection when used in combination with other immunosuppressant drugs
Muromonab-CD3	Monoclonal AB that binds to receptor sites on T-cells, preventing activation by transplanted tissue
Prednisolone	Corticosteroid, suppresses inflammation associated with transplant rejection
Psora-4	Kv1.3-blocker
Rifampicin	Antibiotic; has immunomodulatory properties
Rituximab	CD20 antibody
S100 $\beta$	possesses immunosuppressive activities in diabetic animal models
Sirolimus, Rapamycin	Immunosuppressant and potent inhibitor of cytokine (e.g.IL-2)-dependent T-cell proliferation (kidney transplant)
Tacrolimus (Prograf, FK-506)	Interferes with IL-2 TCR communication
Triptolide	Small molecule; inhibits T-cell activation

The combination therapy may comprise coadministration of the medicaments during the treatment period and/or separate administration of single medicaments during different time intervals in the treatment period.

5 The nucleic acids of the invention or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acids or the proteins are to be assessed. Further antibodies that bind immunospecifically to the novel substances of the invention may be used in therapeutic or diagnostic methods.

10 For example, in one aspect, antibodies, which are specific for a protein of the invention, may be used directly as an effector/modulator, e.g. an antagonist or an agonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express the protein. The  
15 antibodies may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralising antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

20 For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with the protein or any fragment or oligopeptide thereof, which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase  
25 immunological response. It is preferred that the peptides, fragments or oligopeptides used to induce antibodies to the protein have an amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids.

30 Monoclonal antibodies to the proteins may be prepared using any technique that provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Köhler G. and Milstein C., (1975) Nature 256: 495-497; Kozbor D. et al., (1985) J.  
35 Immunol. Methods 81: 31-42; Cote R.J. et al., (1983) Proc. Natl. Acad. Sci. 80: 2026-2030; Cole S.P. et al., (1984) Mol. Cell Biol. 62: 109-120).

In addition, techniques developed for the production of 'chimeric antibodies', the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison S.L. et al., (1984) Proc. Natl. Acad. Sci. 81: 6851-6855; Neuberger M.S. et al., (1984) Nature 312: 604-608; Takeda S. et al., (1985) Nature 314: 452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce single chain antibodies specific for a protein of the invention. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Kang A.S. et al., (1991) Proc. Natl. Acad. Sci. 88: 11120-11123). Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi R. et al., (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; Winter G. and Milstein C., (1991) Nature 349: 293-299).

Antibody fragments, which contain specific binding sites for the proteins may also be generated. For example, such fragments include, but are not limited to, the F(ab')<sub>2</sub> fragments which can be produced by Pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse W.D. et al., (1989) Science 246: 1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding and immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between the protein and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering protein epitopes is preferred, but a competitive binding assay may also be employed (Maddox, supra).

In another embodiment of the invention, the polynucleotides or fragments

thereof or nucleic acid effector/modulator molecules such as antisense molecules, aptamers, RNAi molecules or ribozymes may be used for therapeutic purposes. In one aspect, aptamers, i.e. nucleic acid molecules, which are capable of binding to a protein of the invention and modulating its activity, may be generated by a screening and selection procedure involving the use of combinatorial nucleic acid libraries.

In a further aspect, antisense molecules may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding SF1-SF8 or homologous proteins. Thus, antisense molecules may be used to modulate/effect protein activity or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding the proteins. Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods, which are well known to those skilled in the art, can be used to construct recombinant vectors, which will express antisense molecules complementary to the polynucleotides of the genes encoding the proteins of the invention. These techniques are described both in Sambrook et al. (supra) and in Ausubel et al. (supra). Genes encoding the proteins of the invention can be turned off by transforming a cell or tissue with expression vectors, which express high levels of polynucleotides that encode the proteins of the invention or functional fragments thereof. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, e.g. DNA, RNA or nucleic acid analogues such as PNA, to the control regions of the genes encoding SF1-SF8, i.e., the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start

site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it cause inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors or regulatory molecules. Recent therapeutic  
5 advances using triplex DNA have been described in the literature (Gee J.E. et al., (1994) Gene 149: 109-114; Huber B.E. and Carr B.I., (1994) Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y.). The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

10 Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples, which may be  
15 used, include engineered hammerhead motif ribozyme molecules that can be specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding the proteins of the invention. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA,  
20 GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary  
25 oligonucleotides using ribonuclease protection assays.

Nucleic acid effector/modulator molecules, e.g. antisense molecules and ribozymes may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically  
30 synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences. Such DNA sequences may be incorporated into a variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize  
35 antisense RNA constitutively or inducibly can be introduced into cell lines, cells or tissues. RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition

of flanking sequences at the 5' and/or 3' ends of the molecule or modifications in the nucleobase, sugar and/or phosphate moieties, e.g. the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of non-traditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods, which are well known in the art. Any of the therapeutic methods described above may be applied to any suitable subject including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of SF1-SF8 nucleic acids or the proteins, antibodies to the proteins of the invention, mimetics, agonists, antagonists or inhibitors of the proteins of the invention or nucleic acids. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone or in combination with other agents, drugs or hormones. The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may



contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries, which facilitate processing of the active compounds into preparations, which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.).

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art. For any compounds, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of preadipocyte cell lines or in animal models, usually mice, rabbits, dogs or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. A therapeutically effective dose refers to that amount of active ingredient, for example the SF1-SF8 nucleic acids or proteins or fragments thereof or antibodies, which is sufficient for treating a specific condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions, which exhibit large therapeutic indices, are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration. The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors, which may be taken into account, include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions

may be administered every 3 to 4 days, every week or once every two weeks depending on half-life and clearance rate of the particular formulation. Normal dosage amounts may vary from 0.1 to 100,000 microg, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

In another embodiment, antibodies which specifically bind to the proteins may be used for the diagnosis of conditions or diseases characterized by or associated with over- or under-expression of the proteins of the invention or in assays to monitor patients being treated with the proteins of the invention, or effectors/modulators thereof, e.g. agonists, antagonists, or inhibitors. Diagnostic assays include methods which utilize the antibody and a label to detect the protein in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules, which are known in the art may be used several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring proteins are known in the art and provide a basis for diagnosing altered or abnormal levels of gene expression. Normal or standard values for gene expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibodies to the protein under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, but preferably by photometric means. Quantities of protein expressed in control and disease, samples e.g. from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides specific for the SF1-SF8 proteins may be used for diagnostic purposes. The polynucleotides, which may be used, include oligonucleotide sequences, antisense RNA and

DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which gene expression may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess gene expression, and to monitor regulation of protein levels during therapeutic intervention.

In one aspect, hybridization with probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding the proteins of the invention or closely related molecules, may be used to identify nucleic acid sequences which encode the respective protein. The hybridization probes of the subject invention may be DNA or RNA and are preferably derived from the nucleotide sequence of the polynucleotide encoding the proteins of the invention or from a genomic sequence including promoter, enhancer elements, and introns of the naturally occurring gene. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as  $^{32}\text{P}$  or  $^{35}\text{S}$  or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences specific for SF1-SF8 proteins or homologous nucleic acids may be used for the diagnosis of conditions or diseases, which are associated with the expression of the proteins. Examples of such diseases include the pancreatic diseases (e.g. diabetes), obesity, metabolic syndrome, and/or others. Polynucleotide sequences specific for the SF1-SF8 proteins may also be used to monitor the progress of patients receiving treatment for pancreatic diseases (e.g. diabetes), obesity, and/or metabolic syndrome. The polynucleotide sequences may be used qualitative or quantitative assays, e.g. in Southern or Northern analysis, dot blot or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or chip assays utilizing fluids or tissues from patient biopsies to detect altered gene expression.

In a particular aspect, the SF1-SF8 nucleotide sequences may be useful in assays that detect activation or induction of various metabolic diseases or dysfunctions. The nucleotide sequences may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period,

the sample is washed and the signal is quantitated and compared with a standard value. The presence of altered levels of nucleotide sequences encoding the proteins of the invention in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials or in monitoring the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disease associated with expression of the SF1-SF8 proteins, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence or a fragment thereof, which is specific for the nucleic acids encoding the proteins of the invention, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease. Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that, which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to pancreatic diseases (e.g. diabetes), obesity, and/or metabolic syndrome, the presence of an unusual amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the metabolic diseases and disorders.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding the proteins of the invention may involve the use of PCR. Such

oligomers may be chemically synthesized, generated enzymatically or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5prime.fwdarw.3prime) and another with antisense (3prime.rarw.5prime), employed under optimized  
5 conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantification of closely related DNA or RNA sequences.

10 In another embodiment of the invention, the nucleic acid sequences may also be used to generate hybridization probes, which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. Such techniques include FISH, FACS or artificial  
15 chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price C.M., (1993) Blood Rev. 7: 127-134, and Trask B.J., (1991) Trends Genet. 7: 149-154. FISH (as described in Verma R.S. and Babu A., (1989) Human Chromosomes: A Manual of Basic Techniques,  
20 Pergamon Press, New York, N.Y.). The results may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265: 1981f). Correlation between the location of the gene encoding the proteins of the invention on a physical chromosomal map and a specific disease or  
25 predisposition to a specific disease, may help to delimit the region of DNA associated with that genetic disease.

The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.  
30 An analysis of polymorphisms, e.g. single nucleotide polymorphisms may be carried out. Further, in situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such  
35 as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms or parts thereof, by physical mapping. This provides

valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti R.A. et al., (1988) Nature 336: 577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequences of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

In another embodiment of the invention, the proteins of the invention, their catalytic or immunogenic fragments or oligopeptides thereof, an in vitro model, a genetically altered cell or animal, can be used for screening libraries of compounds in any of a variety of drug screening techniques. One can identify effectors, e.g. receptors, enzymes, proteins, ligands, or substrates that bind to, modulate or mimic the action of one or more of the SF1-SF8 proteins of the invention. The protein or fragment thereof employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between the SF1-SF8 proteins of the invention and the agent tested, may be measured. Agents could also, either directly or indirectly, influence the activity of the proteins of the invention.

In addition activity of the proteins of the invention against their physiological substrate(s) or derivatives thereof could be measured in cell-based or cell-free assays. Agents may also interfere with posttranslational modifications of the protein, such as phosphorylation and dephosphorylation, farnesylation, palmitoylation, acetylation, alkylation, ubiquitination, proteolytic processing, subcellular localization and degradation. Moreover, agents could influence the dimerization or oligomerization of the proteins of the invention or, in a heterologous manner, of the proteins of the invention with other proteins, for example, but not exclusively, docking proteins, enzymes, receptors, or translation factors. Agents could also act on the physical interaction of the proteins of this invention with other proteins, which are required for protein function, for example, but not exclusively, their downstream signaling.

Methods for determining protein-protein interaction are well known in the art.

For example binding of a fluorescently labeled peptide derived from the interacting protein to the SF1-SF8 protein of the invention, or vice versa, could be detected by a change in polarisation. In case that both binding partners, which can be either the full length proteins as well as one binding partner as the full length protein and the other just represented as a peptide are fluorescently labeled, binding could be detected by fluorescence energy transfer (FRET) from one fluorophore to the other. In addition, a variety of commercially available assay principles suitable for detection of protein-protein interaction are well known in the art, for example but not exclusively AlphaScreen (PerkinElmer) or Scintillation Proximity Assays (SPA) by Amersham. Alternatively, the interaction of the SF1-SF8 proteins of the invention with cellular proteins could be the basis for a cell-based screening assay, in which both proteins are fluorescently labeled and interaction of both proteins is detected by analysing cotranslocation of both proteins with a cellular imaging reader, as has been developed for example, but not exclusively, by Cellomics or EvotecOAI. In all cases the two or more binding partners can be different proteins with one being the protein of the invention, or in case of dimerization and/or oligomerization the protein of the invention itself.

Of particular interest are screening assays for agents that have a low toxicity for mammalian cells. The term "agent" as used herein describes any molecule, e.g. protein or pharmaceutical, with the capability of altering or mimicking the physiological function of one or more of the proteins of the invention. Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 Daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise carbocyclic or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups.

Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, nucleic acids and derivatives, structural analogs or combinations thereof. Candidate agents are obtained from a wide variety of sources including libraries of synthetic or

5 natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and oligopeptides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known pharmacological agents may be subjected to directed or random chemical  
10 modifications, such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs. Where the screening assay is a binding assay, one or more of the molecules may be joined to a label, where the label can directly or indirectly provide a detectable signal.

15 Another technique for drug screening, which may be used, provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to the proteins of the invention large numbers of different small test compounds, e.g. aptamers, peptides, low-molecular weight compounds etc., are provided or synthesized on a solid substrate, such as  
20 plastic pins or some other surface. The test compounds are reacted with the proteins or fragments thereof, and washed. Bound proteins are then detected by methods well known in the art. Purified proteins can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide  
25 and immobilize it on a solid support. In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding the protein specifically compete with a test compound for binding the protein. In this manner, the antibodies can be used to detect the presence of  
30 any peptide, which shares one or more antigenic determinants with the protein.

Compounds that bind SF1-SF8 proteins, e.g. antibodies, are useful for the identification or enrichment of cells, which are positive for the expression of the proteins of the invention, from complex cell mixtures. Such cell populations are  
35 useful in transplantation, for experimental evaluation, and as source of lineage and cell specific products, including mRNA species useful in identifying genes specifically expressed in these cells, and as target for the identification of



factors of molecules that can affect them. Cells expressing the protein of the invention or which have been treated with the protein of the invention are useful in transplantation to provide a recipient with pancreatic islet cells, including insulin producing beta cells; for drug screening; experimental models of islet differentiation and interaction with other cell types; in vitro screening assays to define growth and differentiation factors, and to additionally characterize genes involved in islet development and regulation; and the like. The native cells may be used for these purposes, or they may be genetically modified to provide altered capabilities. Cells from a regenerating pancreas, from embryonic foregut, stomach and duodenum, or other sources of pancreatic progenitor cells may be used as a starting population. The progenitor cells may be obtained from any mammalian species, e.g. equine, bovine, porcine, canine, feline, rodent, e.g. mice, rats, hamster, primate, etc. particularly human.

In another embodiment, in a high-throughput screening method, the cells are transfected with a DNA construct, e.g. a viral or non-viral vector containing a reporter gene, e.g. the lacZ gene or the GFP gene, under regulatory control of a promoter of a gene involved in for example beta-cell differentiation, e.g. a promoter of a gene stimulation beta-cell differentiation, preferably a Pax4 promoter. The transfected cells are divided into aliquots and each aliquot is contacted with a test substance, e.g., candidate 1, candidate 2 and candidate 3. The activity of the reporter gene corresponds to the capability of the test compound to induce beta-cell differentiation.

In a further embodiment, which may be combined with the high-throughput screening as described above, a medium throughput validation is carried out. Therein, the test compound is added to stem cells being cultivated and the insulin production is determined. Following an initial high throughput assay, such as the cell based assay outlined above where for example a Pax4 promoter is used as marker for beta-cell regeneration, the activity of candidate molecules to induce beta-cell differentiation is tested in a validation assay comprising adding said compounds to the culture media of the embryoid bodies. Differentiation into insulin-producing cells is then evaluated, e.g. by comparison to wild type and/or Pax4 expressing ES cells to assess the effectiveness of a compound.

The nucleic acids encoding the SF1-SF8 proteins of the invention can be used to generate transgenic cell lines and animals. These transgenic non-human animals are useful in the study of the function and regulation of the proteins of the invention in vivo. Transgenic animals, particularly mammalian transgenic animals, can serve as a model system for the investigation of many developmental and cellular processes common to humans. A variety of non-human models of metabolic disorders can be used to test modulators of the protein of the invention. Misexpression (for example, overexpression or lack of expression) of the protein of the invention, particular feeding conditions, and/or administration of biologically active compounds can create models of metabolic disorders.

In one embodiment of the invention, such assays use mouse models of insulin resistance and/or diabetes, such as mice carrying gene knockouts in the leptin pathway (for example, ob (leptin) or db (leptin receptor) mice), as described above. In addition to testing the expression of the proteins of the invention in such mouse strains (see Examples), these mice could be used to test whether administration of a candidate modulator alters for example lipid accumulation in the liver, in plasma, or adipose tissues using standard assays well known in the art, such as FPLC, colorimetric assays, blood glucose level tests, insulin tolerance tests and others.

Transgenic animals may be made through homologous recombination in embryonic stem cells, where the normal locus of the gene encoding the protein of the invention is mutated. Alternatively, a nucleic acid construct encoding the protein is injected into oocytes and is randomly integrated into the genome. One may also express the genes of the invention or variants thereof in tissues where they are not normally expressed or at abnormal times of development. Furthermore, variants of the genes of the invention like specific constructs expressing anti-sense molecules or expression of dominant negative mutations, which will block or alter the expression of the proteins of the invention may be randomly integrated into the genome. A detectable marker, such as lac Z or luciferase may be introduced into the locus of the genes of the invention, where upregulation of expression of the genes of the invention will result in an easily detectable change in phenotype. Vectors for stable integration include plasmids, retroviruses and other animal viruses, yeast artificial chromosomes (YACs), and the like. DNA constructs for homologous

recombination will contain at least portions of the genes of the invention with the desired genetic modification, and will include regions of homology to the target locus. Conveniently, markers for positive and negative selection are included. DNA constructs for random integration do not need to contain regions of homology to mediate recombination. DNA constructs for random  
5 integration will consist of the nucleic acids encoding the proteins of the invention, a regulatory element (promoter), an intron and a poly-adenylation signal. Methods for generating cells having targeted gene modifications through homologous recombination are known in the field. For embryonic stem  
10 (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, e.g. mouse, rat, guinea pig, etc. Such cells are grown on an appropriate fibroblast-feeder layer and are grown in the presence of leukemia inhibiting factor (LIF). ES or embryonic cells may be transfected and can then be used to produce transgenic animals. After transfection, the ES  
15 cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be selected by employing a selection medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination. Colonies that are positive may then be used for embryo manipulation and morula aggregation. Briefly, morulae are  
20 obtained from 4 to 6 week old superovulated females, the Zona Pellucida is removed and the morulae are put into small depressions of a tissue culture dish. The ES cells are trypsinized, and the modified cells are placed into the depression closely to the morulae. On the following day the aggregates are transferred into the uterine horns of pseudopregnant females. Females are then  
25 allowed to go to term. Chimeric offsprings can be readily detected by a change in coat color and are subsequently screened for the transmission of the mutation into the next generation (F1-generation). Offspring of the F1-generation are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous  
30 progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogenic or congenic grafts or transplants, or in vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animal, domestic animals, etc., for example, mouse, rat, guinea pig, sheep, cow, pig, and others. The transgenic animals  
35 may be used in functional studies, drug screening, and other applications and are useful in the study of the function and regulation of the proteins of the invention in vivo.

Finally, the invention also relates to a kit comprising at least one of

- (a) a nucleic acid molecule coding for a protein of the invention or a functional fragment thereof;
- 5 (b) a protein of the invention or a fragment or an isoform thereof;
- (c) a vector comprising the nucleic acid of (a);
- (d) a host cell comprising the nucleic acid of (a) or the vector of (b);
- (e) a polypeptide encoded by the nucleic acid of (a);
- (f) a fusion polypeptide encoded by the nucleic acid of (a);
- 10 (g) an antibody, an aptamer or another effector / modulator against the nucleic acid of (a) or the polypeptide of (b), (e) or (f) and
- (h) an anti-sense oligonucleotide of the nucleic acid of (a).

15 The kit may be used for diagnostic or therapeutic purposes or for screening applications as described above. The kit may further contain user instructions.

The Figures show:

20 **Fig. 1** shows the analysis of SF1 expression in mammalian (mouse) tissues.

**Fig. 1A** shows the real-time PCR analysis of SF1 expression in wild type mouse tissues (referred to as wt-mice) and in tissues of mice fed with a control diet (referred to as controldiet).

25 **Fig. 1B** shows the real-time PCR analysis of SF1 expression in genetically obese mice (referred to as ob/ob-mice) compared to wild type mice, and in mice fed with a high fat diet (referred to as HFD-mice) compared to mice fed with a control diet.

**Fig. 2** shows the analysis of SF2 expression in mammalian (mouse) tissues.

30 **Fig. 2A** shows the real-time PCR analysis of SF2 expression in wild type mouse tissues (referred to as wt-mice) and in tissues of mice fed with a control diet (referred to as controldiet).

35 **Fig. 2B** shows the real-time PCR analysis of SF2 expression in genetically obese mice (referred to as ob/ob-mice) compared to wild type mice, and in mice fed with a high fat diet (referred to as HFD-mice) compared to mice fed with a control diet.

**Fig. 3** shows the analysis of SF3 expression in mammalian (mouse) tissues.

**Fig. 3A** shows the real-time PCR analysis of SF3 expression in wild type mouse tissues (referred to as wt-mice) and in tissues of mice fed with a control diet (referred to as controldiet).

**Fig. 3B** shows the real-time PCR analysis of SF3 expression in genetically obese mice (referred to as ob/ob-mice) compared to wild-type mice, and in mice fed with a high fat diet (referred to as HFD-mice) compared to mice fed with a control diet.

**Fig. 4** shows the analysis of SF4 expression in mammalian (mouse) tissues.

**Fig. 4A** shows real-time PCR the analysis of SF4 expression in wild type mouse tissues (referred to as wt-mice) and in tissues of mice fed with a control diet (referred to as controldiet).

**Fig. 4B** shows the real-time PCR analysis of SF4 expression in fasted mice (referred to as fasted-mice) and genetically obese mice (referred to as ob/ob-mice) compared to wild-type mice, and in mice fed with a high fat diet (referred to as HFD-mice) compared to mice fed with a control diet.

**Fig. 4C** shows the real-time PCR analysis of SF4 expression in mouse tissues from non-obese-diabetic (NOD) mice compared to wild-type mice.

**Fig. 5** shows the analysis of SF5 expression in mammalian (mouse) tissues.

**Fig. 5A** shows the real-time PCR analysis of SF5 expression in wild type mouse tissues (referred to as wt-mice) and in tissues of mice fed with a control diet (referred to as controldiet).

**Fig. 5B** shows the real-time PCR analysis of SF5 expression in genetically obese mice (referred to as ob/ob-mice) compared to wild type mice, and in mice fed with a high fat diet (referred to as HFD-mice) compared to mice fed with a control diet.

**Fig. 6** shows the analysis of SF6 expression in mammalian (mouse) tissues.

**Fig. 6A** shows the real-time PCR analysis of SF6 expression in wild type mouse tissues (referred to as wt-mice) and in tissues of mice fed with a control diet (referred to as controldiet).

**Fig. 6B** shows the real-time PCR analysis of SF6 expression in tissues of mice fed with a control diet (without brain).

**Fig. 6C** shows the real-time PCR analysis of SF6 expression in genetically obese mice (referred to as ob/ob-mice) compared to wild type mice, and in mice fed with a high fat diet (referred to as HFD-mice) compared to mice fed with a control diet.

5

**Fig. 7** shows the analysis of SF7 expression in mammalian (mouse) tissues.

**Fig. 7A** shows the real-time PCR analysis of SF7 expression in wild type mouse tissues (referred to as wt-mice) and in tissues of mice fed with a control diet (referred to as controldiet).

10

**Fig. 7B** shows the real-time PCR analysis of SF7 expression in genetically obese mice (referred to as ob/ob-mice) compared to wild type mice, and in mice fed with a high fat diet (referred to as HFD-mice) compared to mice fed with a control diet.

15

**Fig. 8** shows the analysis of SF8 expression in mammalian (mouse) tissues.

**Fig. 8A** shows the real-time PCR analysis of SF8 expression in wild type mouse tissues (referred to as wt-mice) and in tissues of mice fed with a control diet (referred to as controldiet).

20

**Fig. 8B** shows the real-time PCR analysis of SF8 expression in genetically obese mice (referred to as ob/ob-mice) compared to wild type mice, and in mice fed with a high fat diet (referred to as HFD-mice) compared to mice fed with a control diet.

25

**Fig. 9** shows the microarray analysis of SF3 expression in human abdominal adipocyte cells during the differentiation from preadipocytes to mature adipocytes.

30

**Fig. 10A** shows the microarray analysis of SF5 expression in human abdominal adipocyte cells during the differentiation from preadipocytes to mature adipocytes.

**Fig. 10B** shows the microarray analysis of SF5 expression in human SGBS cells during the differentiation from preadipocytes to mature adipocytes.

35

**Fig. 11** shows in situ hybridization results for the SF1 protein of the invention.

**Fig. 11A** shows a cryosection of embryonic mouse pancreas (day E17.5),

stained with an in situ hybridization method.

**Fig. 11B** shows a cryosection of embryonic mouse pancreas (day E17.5), stained with an in situ hybridization method and co-stained with the exocrine marker amylase (red).

5 **Fig. 11C** shows a cryosection of embryonic mouse pancreas (day E17.5), stained with an in situ hybridization method and co-stained with the beta cell marker insulin (green).

**Fig. 11D** shows a cryosection of embryonic mouse pancreas (day E17.5), stained with an in situ hybridization method, and co-stained with the exocrine marker amylase (red).

10 **Fig. 11E** shows a cryosection of embryonic mouse pancreas (day E17.5) stained with an in situ hybridization method, and co-stained with the beta cell marker insulin (green).

**Fig. 11F** shows a cryosection of embryonic mouse pancreas (day E17.5), stained with an in situ hybridization method.

15 **Fig. 11G** shows a cryosection of embryonic mouse pancreas (day E17.5) stained with an in situ hybridization method, and co-stained with the beta cell marker insulin (green) and the duct cell marker DBA lectin (blue).

20 The examples illustrate the invention:

### **Example 1: Identification of secreted factors expressed in pancreas**

25 A screen for secreted factors expressed in developing mouse pancreas was carried out according to methods known by those skilled in the art (see, for example Pera E.M. and De Robertis E.M., (2000) Mech Dev 96(2): 183-195) with several modifications.

#### **Expression cDNA library:**

30 During organogenesis, the pancreatic bud is surrounded and influenced by the associated mesenchyme. (see for example, Madsen O.D. et al., (1996) Eur. J. Biochem. 242: 435-445 and Slack, J.M., (1995) Development 121: 1569-1580). Recently, it was suggested, that white adipocytes originate directly from mesenchymal cells (Atanossova P.K., (2003) Folia Med. 45: 41-45). During  
35 embryogenesis, the innervation and vascularization of the pancreas can be observed. Therefore, the tissue used in the screen might have contained

besides pancreatic cells some adipocyte precursors, blood vessels, as well as neuronal cells.

5 A mouse embryonic stage 9.5-15 pancreatic bud library was prepared in pCMVSPORT-6 vector using SUPERScript Plasmid System from Invitrogen according to the manufacturer's instructions. The non-amplified library was electroporated into MaxEff DH10B cells (Invitrogen).

#### Secretion cloning

10 Bacterial clones were picked with sterile toothpicks from agar plates and cultured in 96-deep-well microtiter plates in LB-ampicillin (see Sambrook et al., supra). Aliquots of 8 cultures were pooled, and plasmid DNA was isolated using the BioRobot\_9600 apparatus according to the manufacturer's instructions (Qiagen; QIAprep(r) Turbo BioRobot Kit. Human 293 cell culture  
15 cells were cultured in 75 ml tissue culture flasks in DMEM and 10% fetal calf serum. At 90-99% confluence, the cells were splitted at 1:3 ratio and plated onto poly-D-lysine (Sigma) coated 96-well plates. Cells were transfected with 100-500 ng plasmid using lipofectamine 2000 (Invitrogen). After 6 hours, the medium was exchanged for fresh complete growth medium. 24 hours after  
20 transfection, the cells were washed twice with DMEM without cysteine and methionine (Invitrogen), supplemented with 1% dialysed bovine serum (Sigma) with 50 µg/ml per ml Heparin (Sigma) and glutamine. The cells were labeled radioactively ('S35 Met-label', from Hartmann Analytic GmbH). After 12 hours, aliquots of the supernatants were harvested in 96-well PCR plates and  
25 subjected to SDS gel electrophoresis in precast 4-20% gradient polyacrylamide Criterion gels (Biorad) under reducing conditions, using Criterion Dodeca Cell gel running chamber (Biorad). The gels were fixed in 10% acetic acid, 25% isopropanol for 30 min, soaked 15-30 min in AMPLIFY reagent (Amersham), dried and exposed to X-OMAT (AR) film (Kodak).  
30 Positive clones were identified and regrown in 96-well-plates. DNA of individual clones was prepared and used for transfection as described above. If one of the clones yielded proteins of the same size as that of the original pool, a positive clone was identified. Positive clones were partially sequenced from the 5' end (SEQLAB, Goettingen).



## Example 2: Identification of the human homologous nucleic acid and protein sequences

The term "polynucleotide comprising the nucleotide sequence as shown in GenBank Accession number" relates to the expressible gene of the nucleotide sequences deposited under the corresponding GenBank Accession number. The term "GenBank Accession number" relates to NCBI GenBank database entries (Ref.: Benson D.A. et al., (2000) Nucleic Acids Res. 28: 15-18).

SF1-SF8 homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or vertebrate species, e.g. mammals or birds. Particularly preferred are nucleic acids comprising human homologs as described in Table 1. The mouse sequences identified in the 'secreted factor screen', are also shown in Table 1.

Sequences homologous to the mouse sequences were identified using the publicly available program BLASTP 2.2.3 of the non-redundant protein data base of the National Center for Biotechnology Information (NCBI) (see, Altschul S.F. et al., (1997) Nucleic Acids Res. 25: 3389-3402). The best human homolog of the mouse sequences are shown in Table 2.

**Table 2: Mammalian genes and proteins of the invention (SF1-SF8)**

Name	Genbank Accession Numbers			
	Mus musculus genes and proteins		Homo sapiens genes and proteins	
	cDNA	Protein	cDNA	Protein
SF1	NM_026522	NP_080798	NM_023947	NP_076436
SF2	NM_023625	NP_076114	NM_173542	NP_775813
SF3	AB_051390	BAB18461	NM_145584	NP_663559
SF4	NM_009312	NP_033338	NM_013251	NP_037383
SF5	AB025280	BAA76386	NM_005928	NP_005919
SF6	NM_016743	NP_058023	NM_006159	NP_006150
SF7	NM_007618	NP_031644	NM_001756	NP_001747
SF8	NM_008908	NP_032934	NM_000943	NP_000934

### **Example 3: Taqman analysis of the expression of the nucleic acids of the invention in mammalian (mouse) tissues**

To analyse the expression of the SF1-SF8 mRNAs disclosed in this invention in mammalian tissues, several mouse strains (preferably mice strains C57Bl/6J, C57Bl/6 ob/ob, C57Bl/KS db/db, and Non-Obese-Diabetic (NOD) mice, which are standard model systems in obesity and diabetes research) were purchased from Harlan Winkelmann (33178 Borcheln, Germany) and Taconic M & B (Germantown, NY 12526, U.S.A.), respectively, and maintained under constant temperature (preferably 22°C), 40 per cent humidity and a light / dark cycle of preferably 14 / 10 hours. The mice were fed a standard chow (for example, from ssniff Spezialitäten GmbH, order number ssniff M-Z V1126-000). For the fasting experiment ("fasted wild type mice"), wild type mice were starved for 48 h without food, but only water supplied ad libitum. (see, for example, Schnetzler et al. J Clin Invest 1993 Jul;92(1):272-80, Mizuno et al. Proc Natl Acad Sci U S A 1996 Apr 16;93(8):3434-8). In a further experiment wild-type (wt) mice were fed a control diet (preferably Altromin C1057 mod control, 4.5% crude fat) or high fat diet (preferably Altromin C1057mod. high fat, 23.5% crude fat). Animals were sacrificed at an age of 6 to 8 weeks. The animal tissues were isolated according to standard procedures known to those skilled in the art, snap frozen in liquid nitrogen and stored at -80°C until needed.

RNA was isolated from mouse tissues using Trizol Reagent (for example, from Invitrogen, Karlsruhe, Germany) and further purified with the RNeasy Kit (for example, from Qiagen, Germany) in combination with an DNase-treatment according to the instructions of the manufacturers and as known to those skilled in the art. Total RNA was reverse transcribed (preferably using Superscript II RNaseH<sup>-</sup> Reverse Transcriptase, from Invitrogen, Karlsruhe, Germany) and subjected to Taqman analysis preferably using the Taqman 2xPCR Master Mix (from Applied Biosystems, Weiterstadt, Germany; the Mix contains according to the Manufacturer for example AmpliTaq Gold DNA Polymerase, AmpErase UNG, dNTPs with dUTP, passive reference Rox and optimized buffer components) on a GeneAmp 5700 Sequence Detection System (from Applied Biosystems, Weiterstadt, Germany).

The following prime/probe pairs were used for the TaqMan analysis (GenBank Accession Number NM\_026522 (mouse) for the mouse SF1 sequence):

5 Mouse SF1 forward primer (Seq ID NO:1): 5'- GTT GCC AAG GTC TTT GGG AG -3'; mouse SF1 reverse primer (Seq ID NO:2): 5'- CTC ACG ACC ACG TCT CTT CAG T -3'; mouse SF1 Taqman probe (Seq ID NO:3): (5/6-FAM)- AGT TCA CAC AGA TCT CAC CAG TCT GGT TGC -(5/6-TAMRA).

10 The following prime/probe pairs were used for the TaqMan analysis (GenBank Accession Number NM\_023625 (mouse) for the mouse SF2 sequence):

15 Mouse SF2 forward primer (Seq ID NO:4): 5'- GCC ATG GTC CGG CTC A -3'; mouse SF2 reverse primer (Seq ID NO:5): 5'- CTT TGG GTT GCA GGC TTC A -3'; mouse SF2 Taqman probe (Seq ID NO:6): (5/6-FAM)- TGA TTT CCT CCA TGA CCC TCT GTC ATT GT -(5/6-TAMRA).

20 The following prime/probe pairs were used for the TaqMan analysis (GenBank Accession Number AB051390 (mouse) for the mouse SF3 sequence):

25 Mouse SF3 forward primer (Seq ID NO:7): 5'- CCG GGA GAG ATA CGT GAA GC -3'; mouse SF3 reverse primer (Seq ID NO:8): 5'- TCT CCG TCT CCT CCG TGG -3'; mouse SF3 Taqman probe (Seq ID NO:9): (5/6-FAM)- CCG GAA GAC GGC TCA GTG TGC AT -(5/6-TAMRA).

30 The following prime/probe pairs were used for the TaqMan analysis (GenBank Accession Number NM\_009312 (mouse) for the mouse SF4 sequence):

35 Mouse SF4 forward primer (Seq ID NO:10): 5'- AGC TGC CTC CGT CCC TG -3'; mouse SF4 reverse primer (Seq ID NO:11): 5'- TTC AGC AAT CCT TCC AGA GAG AC-3'; mouse SF4 Taqman probe (Seq ID NO:12): (5/6-FAM)- TTC GGA GAC TCT ACG ACA GCC GCC-(5/6-TAMRA).

The following prime/probe pairs were used for the TaqMan analysis (GenBank Accession Number AB025280 (mouse) for the mouse SF5 sequence):

5 Mouse SF5 forward primer (Seq ID NO:13): 5'- CGG GCC AAG ACA ATG ACA TC -3'; mouse SF5 reverse primer (Seq ID NO:14): 5'- GTC TCA TTG CAC ACA AGG CCT -3'; mouse SF5 Taqman probe (Seq ID NO:15): (5/6-FAM)- ACT GCC TCT GCC CTG AAG GCT TCA -(5/6-TAMRA).

10 The following prime/probe pairs were used for the TaqMan analysis (GenBank Accession Number NM\_016743 (mouse) for the mouse SF6 sequence):

15 Mouse SF6 forward primer (Seq ID NO:16): 5'- TCG CTT CAC CGG GTC TTC -3'; mouse SF6 reverse primer (Seq ID NO:17): 5'- GGC CAT TCT TGC ACT GGC -3'; mouse SF6 Taqman probe (Seq ID NO:18): (5/6-FAM)- ATC AAG CAC GGC ACG GAG TGT ACC C -(5/6-TAMRA).

20 The following prime/probe pairs were used for the TaqMan analysis (GenBank Accession Number NM\_007618 (mouse) for the mouse SF7 sequence):

25 Mouse SF7 forward primer (Seq ID NO:19): 5'- CCA TGC TGC AAC TGG ATG AA -3'; mouse SF7 reverse primer (Seq ID NO:20): 5'- CAG GTG TAC AGG AGG GCC A -3'; mouse SF7 Taqman probe (Seq ID NO:21): (5/6-FAM)- AAT GTG CTG CCT GCT GCC ACC A -(5/6-TAMRA).

30 The following prime/probe pairs were used for the TaqMan analysis (GenBank Accession Number NM\_008908 (mouse) for the mouse SF8 sequence):

35 Mouse SF8 forward primer (Seq ID NO:22): 5'- TGG CTG GGT CAG CAT GG -3'; mouse SF8 reverse primer (Seq ID NO:23): 5'- TCA AGG TGA TGA AGA ACT GGG A -3'; mouse SF8 Taqman probe (Seq ID NO:24): (5/6-FAM)- CAA TGC TGG ACC AGA CAC CAA CGG -(5/6-TAMRA).

In the figures the relative RNA-expression is shown on the Y-axis. In Fig. 1 2, 3, 4, 5, 6, 7, and 8, the tissues tested are given on the X-axis. "WAT" refers to white adipose tissue, "BAT" refers to brown adipose tissue. In Fig. 1 2, 3, 5, 6, 7, and 8, the panel of the wild type mice tissues comprises liver, pancreas, muscle, small intestine, WAT, hypothalamus, and heart, and the panel of the control diet-mice tissues comprises liver, muscle, small intestine, WAT, brain, and heart. In Fig. 4 the panel of the wild type mice tissues comprises WAT, BAT, muscle, liver, pancreas, hypothalamus, brain, testis, colon, small intestine, heart, lung, spleen, and kidney, and the panel of the control diet-mice tissues comprises WAT, BAT, muscle, liver, brain, testis, colon, small intestine, heart, lung, spleen, and kidney.

The function of the proteins of the invention in metabolism was further validated by analyzing the expression of the transcripts in different tissues. For this passage, mouse models of insulin resistance and/or diabetes were used, such as mice carrying gene knockouts in the leptin pathway (for example, ob/ob (leptin) or db/db (leptin receptor/ligand) mice) to study the expression of the proteins of the invention. Such mice develop typical symptoms of diabetes, show hepatic lipid accumulation and frequently have increased plasma lipid levels (see Bruning J.C. et al, (1998) Mol. Cell. 2: 559-569).

Further, expression of the mRNAs encoding the proteins of the invention was also examined in susceptible wild type mice (for example, C57Bl/6) that show symptoms of diabetes, lipid accumulation, and high plasma lipid levels, if fed a high fat diet.

Expression profiling studies confirm the particular relevance of the proteins of the present invention as regulators of energy metabolism in mammals.

Taqman analysis revealed that SF1 is expressed in several mammalian tissues, showing highest level of expression in white adipose tissue (WAT) and brain, and lower but still robust levels in further tissues, e.g. hypothalamus, small intestine, heart, liver, muscle, and pancreas, as shown in Fig. 1A. We found, for example, that the expression of SF1 is down-regulated in the liver, pancreas, muscle, and heart of ob/ob mice compared to wild-type mice (see Fig. 1B). In wild type mice fed a high fat diet, the expression of SF1 is not regulated. The ubiquitous and robust expression levels of SF1 suggest that it

plays an essential role in cellular metabolism. The regulation of gene expression in the mouse model for the metabolic syndrome as described above suggests that it also might play a role in the regulation of energy homeostasis.

5

Taqman analysis revealed that SF2 is expressed in several mammalian tissues, showing highest level of expression in WAT, brain, and hypothalamus, and higher levels in further tissues, e.g. liver, heart, and muscle. Furthermore SF2 is expressed on lower but still robust levels in the small intestine and pancreas as depicted in Fig. 2A. We found, for example, that the expression of SF2 is slightly down-regulated in the pancreas of ob/ob mice compared to wild-type mice (see Fig. 2B). In wild type mice fed a high fat diet, the expression of SF2 is not regulated. The ubiquitous and robust expressssion levels of SF2 suggest that it plays an essential role in cellular metabolism. The regulation of gene expression in the mouse model for the metabolic syndrome as described above suggests that it also might play a role in the regulation of energy homeostasis.

10

15

Taqman analysis revealed that SF3 is expressed in several mammalian tissues, showing highest level of expression in brain and hypothalamus, and higher levels in further tissues, e.g. WAT and heart. Furthermore SF3 is expressed on lower but still robust levels in the muscle and small intestine as depicted in Fig. 3A. Even though SF3 in pancreas of wild type mice was expressed on a low level, a clear down-regulation could be observed in the pancreas of ob/ob-mice (Fig. 3B). In wild type mice fed a high fat diet, the expression of SF3 was not regulated. The expressssion of SF3 in metabolic active tissues of wild type and control mice as well as the regulation of gene expression in the mouse model for the metabolic syndrome as described above suggests that it might play a role in the regulation of energy homeostasis.

20

25

30

Taqman analysis revealed that the expression of SF4 is restricted to a few mammalian tissues, with highest expression levels in hypothalamus and brain of wild type mice. In addition, SF4 is expressed in spleen, white adipose tissue (WAT) and at lower levels in small intestine of wild type and control mice as depicted in Fig. 4A. We found, for example, that the expression of SF4 is significantly downregulated in metabolic active tissue (WAT) and upregulated

35

in muscle of genetically induced obese mice (*ob/ob*) compared to wild type mice. Further, we found that the expression of SF4 is significantly downregulated in spleen of starved wild type mice compared to wild type mice fed a standard diet (see Fig. 4B). In wild type mice fed a high fat diet, the expression of SF4 is downregulated in WAT and small intestine, and upregulated in colon, spleen and brown adipose tissue (BAT), compared to mice fed a control diet, supporting that SF4 is involved in the regulation of mammalian metabolism (see Fig. 4B). In one embodiment of the invention, expression of SF4 mRNA was also examined Non-Obese-Diabetic (NOD) mice. In those mice, the expression of SF4 is significantly downregulated in WAT, and upregulated in spleen, lung, muscle, liver, and testis, supporting that SF4 is involved in the regulation of mammalian metabolism (see Fig. 4C).

Taqman analysis revealed that SF5 is highly expressed in WAT. Furthermore SF5 is expressed on lower but still robust levels in the heart, brain, hypothalamus, small intestine, and muscle as depicted in Fig. 5A. Even though SF5 was expressed in the liver and pancreas of wild type mice on a low level, a clear up-regulation could be observed in the liver, and a down-regulation could be observed in the pancreas of *ob/ob*-mice. In wild type mice fed a high fat diet, the expression of SF5 was up-regulated in the WAT, when compared to mice fed a control diet (see Fig. 5B). The high expression level of SF5 in white adipose tissue of wild type and control mice as well as the regulation of gene expression in the mouse models for the metabolic syndrome as described above suggests that it might play a role in the regulation of energy homeostasis.

Taqman analysis revealed that SF6 is highly expressed in the hypothalamus and brain in wild type and control mice as shown in Fig. 6A. Furthermore SF6 is expressed on lower but still robust levels in the heart, small intestine, and WAT of control mice, as depicted in Fig. 6B. Even though SF6 was expressed in the heart and WAT of control mice on a low level, a down-regulation could be observed in the heart of mice fed a high fat diet and in the WAT of *ob/ob* mice (see Fig. 6C). The high expression level of SF6 in the hypothalamus and brain as well as the regulation of gene expression in mouse models for the metabolic syndrome as described above suggests that it might play a role in the regulation of energy homeostasis.

Taqman analysis revealed that SF7 is highly expressed in the liver of wild type and control diet mice. Furthermore SF7 is expressed on lower but still robust levels in the muscle, WAT, and heart as depicted in Fig. 7A. Even though SF7 was expressed in the WAT and heart of wild type mice on a low level, a down-regulation could be observed in the in the WAT and heart of ob/ob-mice. In wild type mice fed a high fat diet, the expression of SF7 was down-regulated in the muscle and heart, when compared to mice fed a control diet (see Fig. 7B). SF7 is expressed on a very low level in the small intestine of wild type mice but up-regulated to robust expression levels in the small intestine of ob/ob mice and HFD mice. The high expression level of SF7 in the liver of wild type and control mice as well as the regulation of gene expression in the mouse models for the metabolic syndrome as described above suggests that it might play a role in the regulation of energy homeostasis.

Taqman analysis revealed that SF8 is expressed in several mammalian tissues, showing highest level of expression in WAT, muscle, small intestine, heart, and pancreas. Furthermore SF8 is expressed on lower but still robust levels in the liver, hypothalamus, and brain as depicted in Fig. 8A. We found, for example, that the expression of SF8 is down-regulated in the pancreas and small intestine, and up-regulated in the WAT and hypothalamus of ob/ob mice compared to wild type mice (see Fig. 8B). In wild type mice fed a high fat diet, the expression of SF8 was not regulated. The expression of SF8 in metabolic active tissues of wild type and control mice as well as the regulation of gene expression in a mouse model for the metabolic syndrome as described above suggests that it might play a role in the regulation of energy homeostasis.

#### **Example 4: Analysis of the differential expression of transcripts of the proteins of the invention in human tissues**

RNA preparation from human primary adipose tissues was done as described in Example 3. The target preparation, hybridization, and scanning was performed as described in the manufactures manual (see Affymetrix Technical Manual, 2002, obtained from Affmetrix, Santa Clara, USA).

The expression analysis (using Affymetrix GeneChips) of the genes using primary human abdominal adipocyte differentiation clearly shows differential



expression of the human SF3 and SF5 genes in adipocytes. Several independent experiments were done. In Fig. 9 and Fig. 10, the fluorescent intensity is shown on the Y-axis, and the X-axis represents the time axis. "d 0" refers to day 0 (start of the experiment), "d12" refers to day 12 of adipocyte differentiation.

The experiments show that the SF3 transcript is most abundant at day 12 compared to day 0 during differentiation (see Fig. 9). The experiments also show that the SF5 transcript is most abundant at day 0 compared to day 12 during differentiation (see Fig. 10A and Fig. 10B).

Thus, the SF3 protein has to be increased in order for the preadipocytes to differentiate into mature adipocytes, and the SF5 protein has to be decreased in order for the preadipocytes to differentiate into mature adipocytes. The SF3 protein in preadipocytes has the potential to enhance adipose differentiation, and the SF5 protein in preadipocytes has the potential to inhibit adipose differentiation. Therefore, the SF3 and SF5 proteins might play an essential role in the regulation of human metabolism, in particular in the regulation of adipogenesis and thus it might be an essential role in pancreatic diseases (e.g. diabetes), obesity, and/or metabolic syndrome.

#### **Example 5: In situ hybridisations**

SF1 is a protein which was isolated in a screen for secreted proteins expressed in the developing mouse pancreas during embryogenesis. The isolation of the corresponding cDNA from an embryonic mouse pancreas library demonstrates that the gene is expressed in this tissue at this time. To gain more information about the expression of this gene, in situ hybridization was performed on sections of embryonic mouse pancreas (day E17.5) according to standard protocols as known to those skilled in the art and as described previously (for example, Pelton, R.W. et al., (1990) Development 110: 609-620; Belo, J.A. et al., (1997) Mech. Dev. 68: 45-57), combined with antibody stainings for insulin (beta cell marker) or DBA lectin (duct cell marker).

SF1 showed widespread expression in non-beta cell and non-duct cells. The

- 50 -

distribution and shape of the expressing cells as well as co-staining with the exocrine marker amylase indicate that SF1 is expressed predominantly in exocrine cells in the developing pancreas.

- 5 The formation of beta cells during embryogenesis or during regenerative processes after birth is a tightly controlled process. The signalling molecules involved in the control of this process are not known completely. It has been shown that the formation of novel beta cells is regulated by signals produced by other cell types in the pancreas, for example endothelial cells (Lammert
- 10 E, et al., (2001) Science 294: 564-567). Thus, SF1 is a candidate regulator of the formation of pancreatic endocrine and other pancreatic cell types. It is noteworthy that despite their homology to chitinases, homologs of SF1 act as growth factors in *Drosophila* (Kawamura et al., supra).

15

For the purpose of the present invention, it will be understood by the person having average skill in the art, that any combination of any feature mentioned throughout the specification is explicitly disclosed herewith.